Maximum simulated likelihood estimation of the seemingly unrelated stochastic frontier regressions

Hung-pin Lai[1]

Abstract

In this paper, we use the maximum simulated likelihood (MSL) approach to estimate multiple stochastic frontier (SF) models with random effects and correlated composite errors. We show that the separate estimation of the single equation ignores the correlation between the composite errors and causes significant efficiency loss in estimation. In addition to using Monte Carlo simulation to investigate the finite sample performance of the simulated estimator, we demonstrate the usefulness of our approach in estimating the technical efficiency of Taiwan's international hotels based on their accommodation and restaurant divisions.

1 Introduction

Since the pioneering work of Aigner et al. (1977), stochastic frontier (SF) analysis has been widely used in productivity and efficiency studies to describe and estimate models of the production frontier. The empirical model typically assumes that a decision-making unit (DMU) employs a single production process or technology in the production of an output using multi-inputs. However, if an organization (a DMU) operates multiple production divisions (sub-DMUs), with each division supported by its own set of resource inputs, these sub-DMUs may be subject to the same random

---

✉ Hung-pin Lai
ecdhpl@ccu.edu.tw

1   Department of Economics, National Chung Cheng University, Chiayi, Taiwan

shocks as the parent DMU. Given that these divisions would share some commonly observed or unobserved characteristics of the parent DMU, the divisions' technical efficiencies may well be correlated. A system of multiple stochastic frontier regressions on the sub-DMUs will be a more appropriate model to investigate a DMU's operation and performance. Since the system estimation takes into account the mutual dependency among the composite errors, the estimator is more efficient than that from the regression-by-regression estimation.

Previously, Lai and Huang (2013) discussed the estimation of a system of stochastic frontier models for cross-section data. When sub-DMUs are observed over time, some unobserved heterogeneity may exist and a model that can capture the panel characteristics will be able to provide more effective estimation and also a better prediction of the inefficiency. The unobserved heterogeneity for the firm level panel data can be incorporated into the model in several ways. For instance, one can introduce the heterogeneity through fixed effects, random effects or heterogeneous variance in the symmetric/one-sided random component. In the model considered in this paper, we assume the firm heterogeneity comes from the random effects as well as heterogeneous variance of the inefficiency. The assumption of random effects can be easily extended to the fixed effects under Mundlak (1978)'s assumption.

There are several applications of the joint estimation of the system of equations. For instance, Huang et al. (2018b) consider a model with two equations, one for cost efficiency and one for market power, where each equation has a composed error. Similarly, Genius et al. (2012) use a system of input demands with composite errors. Other empirical examples include Huang et al. (2017a, b), and Huang et al. (2018a). See also Amsler and Schmidt (2021) for a systematic review. The common feature of these empirical studies is that they all use the cross-sectional approach even though some of them employ panel data. Accordingly, the main objective of this paper is to extend the model of Lai and Huang (2013) so that the model can capture the unobserved panel characteristics of the multiple SF regression models.

The plan of the paper is as follows. In Sect. 2, we introduce a system of panel stochastic frontier regressions, where the composite errors are correlated. Section 3 discusses how to use the simulated maximum likelihood approach to estimate the model. We then discuss how to estimate the inefficiency using the simulated approach in Sect. 4. We examine the finite sample performance of the proposed estimator by Monte Carlo simulation in Sect. 5. An empirical study of the hotel industry in Taiwan is given in Sect. 6, and a summary conclusion is given in Sect. 7.

2 The seemingly unrelated SF panel model

Consider the following production frontiers of J sub-DMUs:

$$
\begin{aligned}
y_{it}^1 &= \beta_0^1 + x_{it}^{1\prime}\beta_1 + \alpha_i^1 + v_{it}^1 - u_{it}^1, \\
y_{it}^2 &= \beta_0^2 + x_{it}^{2\prime}\beta_2 + \alpha_i^2 + v_{it}^2 - u_{it}^2, \\
&\ \vdots
\end{aligned}
$$

$$y_{it}^J = \beta_0^J + x_{it}^{J\prime} \beta_J + \alpha_i^J + v_{it}^J - u_{it}^J, \tag{1}$$

where $j = 1, \ldots, J$, $i = 1, \ldots, N$, and $t = 1, \ldots, T$. y_{it}^j and x_{it}^j are the log output and the log inputs of the jth sub-DMU. In order to implement the maximum likelihood approach to estimate the model, we make the following assumptions about the random components:

[A1]: $\alpha_i^j \sim N(0, \sigma_{\alpha j}^2)$ is the firm-specific random effect. For a fixed i, α_i^j and $\alpha_i^{j\prime}$ are independent to each other for $j \neq j'$.

[A2]: $v_{it}^j \sim N(0, \sigma_{vj}^2)$ is a two-sided symmetric random noise. For fixed i and j, v_{it}^j and $v_{is}^{j\prime}$ are independent across time. For fixed i and t, v_{it}^j and $v_{it}^{j\prime}$ are correlated for $j \neq j'$.

[A3]: $u_{it}^j \sim N^+(0, \sigma_{uj,it}^2)$ is a one-sided random component that captures the inefficiency and its variance. Moreover, $\sigma_{uj,it}$ can be parametrized as $\sigma_{uj,it} = \exp\left(\delta_j' w_{it}^j\right)$, where w_{it}^j is the vector of exogenous determinants of inefficiency u_{it}^j. For fixed i and j, u_{it}^j and $u_{is}^{j\prime}$ are independent across time. For fixed i and t, u_{it}^j and $u_{it}^{j\prime}$ are correlated for $j \neq j'$.

[A4]: The three random components α_i^j, v_{it}^j and u_{it}^j are independent to each other and uncorrelated with x_{it} for fixed i, t and j.

Although we assume the random effects across divisions are independent in assumption [A1], we will discuss in Sect. 3.2 that the assumption can be further extended to include correlated random effects. Assumption [A2] suggests that the v_{it}^j's of different divisions at the same time are correlated with each other. The same scenario applies to, the u_{it}^j's by assumption [A3]. Now, let us define $e_{it}^j = \alpha_i^j + v_{it}^j - u_{it}^j$ as the composite error of the jth equation, then it follows from [A1]–[A3] that e_{it}^j and $e_{it}^{j\prime}$ will be correlated. For simplicity, we also define $\varepsilon_{it}^j = v_{it}^j - u_{it}^j$, which is the composite error containing all time-variant random components, so that $e_{it}^j = \alpha_i^j + \varepsilon_{it}^j$. There are two main distinctions between the model given in (1) and the SF model discussed in Lai and Huang (2013). One is that we extend the cross-section model to the panel model with random firm effects, and the other is that we allow for heteroscedasticity in the distribution of inefficiency.

3 Copulas and the simulated likelihood function

To motivate the simulated likelihood approach, we first discuss the model without random firm effects and then discuss the model with random effects.

3.1 The model without random effects

Consider the special case of the system of equations (1), where $\alpha_i^j = 0$ for all i and j. Let $\theta_j = (\beta_0^j, \beta_j', \sigma_{vj}^2, \delta_j')'$ be a vector of parameters in the jth SF regression and

$\varepsilon_{it}^j = v_{it}^j - u_{it}^j$ denote the composite error of the jth SF regression. In this model, the only source of heterogeneity is the heteroscedastic variance of u_{it}^j and the correlation between equations comes from v_{it}^j and $v_{it}^{j'}$ and/or u_{it}^j and $u_{it}^{j'}$.

Let $F_{\varepsilon^j}(\varepsilon_{it}^j; \theta_j)$ and $f_{\varepsilon^j}(\varepsilon_{it}^j)$ denote the cumulative distribution function (cdf) and probability density function (pdf) of ε_{it}^j, respectively. When the distribution function of u is half normal, i.e., $N^+(0, \sigma_{uj}^2)$, the pdf of the associated composite errors ε_{it}^j is

$$f_{\varepsilon^j}(\varepsilon_{it}^j) = \frac{2}{\sigma_j} \phi\left(\frac{\varepsilon_{it}^j}{\sigma_j}\right) \Phi\left(-\frac{\lambda_j}{\sigma_j}\varepsilon_{it}^j\right), \tag{2}$$

where $\phi(\cdot)$ represents a standard normal density function, $\sigma_j = \sqrt{\sigma_{vj}^2 + \sigma_{uj,it}^2}$ and $\lambda_j = \sigma_{uj,it}^2/\sigma_{vj}$.[1] It can be shown that ε_{it}^j follows a closed skew normal distribution, i.e.,

$$\varepsilon_{it}^j \sim \mathrm{CSN}_{1,1}\left(0, \sigma_j^2, -\frac{\lambda_j^2}{1+\lambda_j^2}, 0, \frac{\lambda_j^2 \sigma_j^2}{\left(1+\lambda_j^2\right)^2}\right), \tag{3}$$

which has the cdf

$$F_{\varepsilon^j}(\varepsilon_{it}^j) = 2 \cdot \Phi_2\left(\begin{pmatrix}\varepsilon_{it}^j \\ 0\end{pmatrix}; \begin{pmatrix}0 \\ 0\end{pmatrix}, \begin{pmatrix}\sigma_j^2 & \frac{\lambda_j^2 \sigma_j^2}{1+\lambda_j^2} \\ \frac{\lambda_j^2 \sigma_j^2}{1+\lambda_j^2} & \frac{\lambda_j^2 \sigma_j^2}{1+\lambda_j^2}\end{pmatrix}\right), \tag{4}$$

where $\Phi_2(\cdot; \mu, \Sigma)$ denotes the cdf of the bivariate normal distribution with mean μ and variance Σ.

For simplicity, we let $\varepsilon_{it} = \left(\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J\right)'$ be a $J \times 1$ vector and $\varepsilon_{i.} = \left(\varepsilon_{i1}', \ldots, \varepsilon_{iT}'\right)'$ be a $JT \times 1$ vector. According to the result of Sklar's theorem (1959) and Schweizer and Sklar (1983), the joint cdf of ε_{it} can be represented as a copula function of its own one-dimensional margins. More specifically,

$$F_\varepsilon\left(\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J\right) = C\left(F_{\varepsilon^1}\left(\varepsilon_{it}^1\right), \ldots, F_{\varepsilon^J}\left(\varepsilon_{it}^J\right); R\right), \tag{5}$$

where R is the vector of parameters of the copula function. The dependence of the marginal distributions is captured by the copula function. The corresponding copula density is

$$c\left(F_{\varepsilon^1}\left(\varepsilon_{it}^1\right), \ldots, F_{\varepsilon^J}\left(\varepsilon_{it}^J\right); R\right) = \frac{\partial^J C\left(F_{\varepsilon^1}\left(\varepsilon_{it}^1\right), \ldots, F_{\varepsilon^J}\left(\varepsilon_{it}^J\right); R\right)}{\partial F_{\varepsilon^1}\left(\varepsilon_{it}^1\right) \ldots \partial F_{\varepsilon^J}\left(\varepsilon_{it}^J\right)}.$$

---

[1] With slight abuse of notation, we abbreviate $\sigma_{j,it}$ and $\lambda_{j,it}$ as σ_j and λ_j to keep the notation simple.

Therefore, it follows from (8) that the joint pdf of ε_{it} can be represented as

$$f_\varepsilon\left(\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J\right) = c\left(F_{\varepsilon^1}\left(\varepsilon_{it}^1\right), \ldots, F_{\varepsilon^J}\left(\varepsilon_{it}^J\right); R\right) \cdot \prod\nolimits_{j=1}^J f_{\varepsilon^J}\left(\varepsilon_{it}^J\right). \quad (6)$$

One special copula is the independent (or product) copula, which is defined as

$$C\left(\varsigma_{it}^1, \ldots, \varsigma_{it}^J\right) = \prod\nolimits_{j=1}^J \varsigma_{it}^j,$$

where $\varsigma_{it}^j = F_{\varepsilon^j}\left(\varepsilon_{it}^j\right)$. The corresponding copula density is

$$c\left(\varsigma_{it}^1, \ldots, \varsigma_{it}^J\right) = 1.$$

Therefore, Eq. (6) suggests that $f_\varepsilon\left(\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J\right) = \prod_{j=1}^J f_{\varepsilon^j}\left(\varepsilon_{it}^j\right)$ when the independent copula is used. It implies $\varepsilon_{it}^{j'}s$ are independent to each other.

Moreover, if the Gaussian copula is assumed, then

$$C\left(F_{\varepsilon^1}\left(\varepsilon_{it}^1\right), \ldots, F_{\varepsilon^J}\left(\varepsilon_{it}^J\right); R\right) = \Phi_R\left(\Phi^{-1}\left(F_{\varepsilon^1}\left(\varepsilon_{it}^1\right)\right), \ldots, \Phi^{-1}\left(F_{\varepsilon^J}\left(\varepsilon_{it}^J\right)\right)\right),$$

where $\Phi_R(\cdot)$ denotes the standardized multivariate normal distribution with the correlation matrix R. The corresponding copula density[2] is

$$c\left(\varsigma_{it}^1, \ldots, \varsigma_{it}^J\right) = \frac{1}{|R|^{1/2}} \exp\left(-\frac{1}{2}\varsigma_{i.}'\left(R^{-1} - I\right)\varsigma_{i.}\right), \quad (7)$$

where $\varsigma_{i.} = (\varsigma_{it}^1, \ldots, \varsigma_{it}^J)'$ and $\varsigma_{it}^j = \Phi^{-1}\left(F_{\varepsilon^j}\left(\varepsilon_{it}^j\right)\right)$. Therefore, the joint pdf of $\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J$ is

$$f_\varepsilon\left(\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J\right) = f_\varepsilon\left(\varepsilon_{it}\right) = c\left(\Phi^{-1}\left(F_{\varepsilon^1}\left(\varepsilon_{it}^1\right)\right), \ldots, \Phi^{-1}\left(F_{\varepsilon^J}\left(\varepsilon_{it}^J\right)\right); R\right)$$
$$\times \prod\nolimits_{j=1}^J f_{\varepsilon^j}\left(\varepsilon_{it}^j\right). \quad (8)$$

The log-likelihood function is

$$\ln L(\theta) = \sum_{i=1}^N \sum_{t=1}^T \ln f_\varepsilon\left(\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J\right). \quad (9)$$

---

[2] For instance, see Cherubini et al. (2004) and Trivedi and Zimmer (2005) for the Gaussian copula density and other possible choices of copula functions.

Let $\theta = (\theta_1', \ldots, \theta_J')'$ and Θ be the parameter space, then the maximum likelihood (ML) estimator of θ is defined as

$$\widehat{\theta}_{\mathrm{ML}} = \arg \underset{\theta \in \Theta}{\mathrm{Max}} \ln L(\theta). \tag{10}$$

Since there are no random firm effects, this model is a pooled system of regressions. Empirically, the true copula is in general unknown, so the sandwich formula is suggested to used for computing the standard errors of the ML estimator.

Note that the current model has a very similar setting with the model discussed in Amsler et al. (2014, APS hereafter). The main difference between them is that the APS model has a single equation, i.e., $J = 1$, which is a special case of our model. Moreover, they use the copula function to model time dependence instead of the correlation between the multi-equations for the J divisions of firm i at time t. Although in principal one can always estimate the model in (1) using equation-by-equation estimation, the estimators of the model parameters are consistent but inefficient due to ignorance of the correlation across equations. Empirically, it is not clear how the random components, such as v_{it}^j and $v_{it}^{j'}$ and/or u_{it}^j and $u_{it}^{j'}$, are correlated. With the copula framework, we can focus on modeling the marginal distributions of each single equation and the dependence between the marginal distributions of the composite errors can be captured by a copula function. Compared with other existing approaches, such as directly specifying the joint distribution of the composite errors, using a copula is relatively easy from the practical point of view.

3.2 The model with random effects

In this section, we discuss estimation of the system of equations (1) with random firm effects. Let $\theta_j = (\beta_0^j, \beta_j', \sigma_{\alpha j}^2, \sigma_{vj}^2, \delta_j')'$ denote the vector of the parameters in the jth division. In the system, the outputs $y_{i1}^j, y_{i2}^j, \ldots, y_{iT}^j$ within the same division are correlated due to their common component α_i^j. Under assumption [A1], the α_i^j's are independent to each other.

To take into account the correlation between the J divisions, let $\alpha_i = (\alpha_i^1, \ldots, \alpha_i^J)'$ be the $J \times 1$ vector of random effects and let $\varepsilon_{it} = (\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J)'$ be defined as before. Recall that $e_{it}^j = \alpha_i^j + v_{it}^j - u_{it}^j$, so we denote $e_{it} = (e_{it}^1, \ldots, e_{it}^J)'$ as a $J \times 1$ vector and $e_{i\cdot} = (e_{i1}', \ldots, e_{iT}')'$ as a $JT \times 1$ vector. The joint pdf of e_{it} in the system SF regression can be evaluated using the simulated approach.[3] It is worth mentioning that $e_{i1}^j, \ldots, e_{iT}^j$ are correlated with each other due to the common component α_i^j, but they are conditionally independent if α_i^j is known. In other words,

$$f_e(e_{it}) = \int f_{e|\alpha}(e_{it}^1, \ldots, e_{it}^J | \alpha_i) f_\alpha(\alpha_i) \mathrm{d}\alpha_i, \tag{11}$$

where $f_{e|\alpha}(e_{it}^1, \ldots, e_{it}^J | \alpha_i) = f_\varepsilon(\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J)$.

---

[3] For instance, see Greene (2003).

The above result suggests that $f_e(e_{it})$ can be evaluated by the simulated joint pdf

$$f_e^s(e_{it}) = \frac{1}{M}\sum_{m=1}^{M} f_{\varepsilon}(\varepsilon_{it(m)}), \tag{12}$$

where $\varepsilon_{it(m)} = \left(\varepsilon_{it(m)}^1, \ldots, \varepsilon_{it(m)}^J\right)'$ is a $J \times 1$ vector, $\varepsilon_{it(m)}^j = e_{it}^j - \alpha_{i(m)}^j$ and $\alpha_{i(m)}^j$ denotes the mth draw from the distribution of α_i^j. The superscript s of $f_e^s(e_{it})$ denotes the simulated density. Therefore, it follows from Eqs. (8) and (12) that the simulated joint pdf $f_e^s(e_{it})$ can be evaluated by

$$f_e^s(e_{it}) = \frac{1}{M}\sum_{m=1}^{M}\left[c\left(F_{\varepsilon^1}\left(\varepsilon_{it(m)}^1\right), \ldots, F_{\varepsilon^J}\left(\varepsilon_{it(m)}^J\right); R\right) \times \prod_{j=1}^{J} f_{\varepsilon^j}^s\left(\varepsilon_{it(m)}^j\right)\right]. \tag{13}$$

For the special case when $J = 1$, Eq. (11) degenerates to the single equation case

$$f_{e^j}(e_{it}^j) = \int f_{e^j|\alpha^j}(e_{it}^j|\alpha_i^j) f_{\alpha^j}(\alpha_i^j) d\alpha_i^j, \tag{14}$$

which can be evaluated by the simulated density

$$f_{e^j}^s(e_{it}^j) = \frac{1}{M}\sum_{m=1}^{M} f_{\varepsilon^j}(e_{it}^j - \alpha_{i(m)}^j) \tag{15}$$

$$= \frac{1}{M}\sum_{m=1}^{M} f_{\varepsilon^j}(\varepsilon_{it(m)}^j), \tag{16}$$

where $\alpha_{i(m)}^j$ denotes the mth Halton draw from α_i^{j}'s distribution $N(0, \sigma_{\alpha j}^2)$.

Now, let us consider the $JT \times 1$ random vector $e_{i.} = \left(e_{i1}', \ldots, e_{iT}'\right)'$. The joint pdf of $e_{i.}$ conditional on α_i is

$$\begin{aligned}
f_{e|\alpha}(e_{i.}|\alpha_i) &= f_{e|\alpha}\left(e_{i1}, \ldots, e_{iT}|\alpha_i\right) \\
&= f_{\varepsilon}(\varepsilon_{i.}) \\
&= \prod_{t=1}^{T} f_{\varepsilon}(\varepsilon_{it}) \\
&= \prod_{t=1}^{T} f_{\varepsilon}\left(\varepsilon_{it}^1, \ldots, \varepsilon_{it}^J\right).
\end{aligned}$$

The third equality is due to that $e_{i1}, \ldots, e_{iT}$ are conditionally independent given α_i. It follows from (13) that the joint pdf of $e_{i.}$ can be approximated by the simulated density

$$\begin{aligned}
f_e^s(e_{i.}) &= \int f_{e|\alpha}^s(e_{i.}|\alpha_i) f_{\alpha}(\alpha_i) d\alpha_i \\
&= \frac{1}{M}\sum_{m=1}^{M}\left[\prod_{t=1}^{T} f_{\varepsilon}(\varepsilon_{it(m)}^1, \ldots, \varepsilon_{it(m)}^J)\right]
\end{aligned}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \prod_{t=1}^{T} \left[c \left(F_{\varepsilon^1} \left(\varepsilon_{it(m)}^1 \right), \ldots, F_{\varepsilon^J} \left(\varepsilon_{it(m)}^J \right); R \right) \times \prod_{j=1}^{J} f_{\varepsilon^j}^s \left(\varepsilon_{it(m)}^j \right) \right].$$

$$(17)$$

The logarithm of the simulated likelihood function of the SF system is

$$\ln L^s \left(\theta_1, \ldots, \theta_J, R \right)$$
$$= \sum_{i=1}^{N} \ln f_e^s (e_{i.})$$
$$= \sum_{i=1}^{N} \ln \left\{ \frac{1}{M} \sum_{m=1}^{M} \prod_{t=1}^{T} \left[\begin{array}{c} c \left(F_{\varepsilon^1} \left(\varepsilon_{it(m)}^1 \right), \ldots, F_{\varepsilon^J} \left(\varepsilon_{it(m)}^J \right); R \right) \\ \times \prod_{j=1}^{J} f_{\varepsilon^j}^s \left(\varepsilon_{it(m)}^j \right) \end{array} \right] \right\}.$$

$$(18)$$

Let $\theta = \left(\theta_1', \ldots, \theta_J', R \right)'$ be the set of all parameters and let Θ denote the parameter space, then the maximum simulated likelihood (MSL) estimator of θ is defined as

$$\widehat{\theta}_{\mathrm{MSL}} = \arg \operatorname*{Max}_{\theta \in \Theta} \ln L^s \left(\theta \right). \tag{19}$$

Similar to the ML estimator in Sect. 3.1, the sandwich formula is suggested to used for computing the standard errors of the MSL estimator.

Note that in the above discussion we assumes the random effects α_i^j's are independent to each other. In other words, in the MSL estimation we assume $f_\alpha(\alpha_i) = \prod_{j=1}^{J} f_{\alpha^j}(\alpha_i^j)$ in Eq. (11) and draw α_i^j for each j and i independently. We may generalize our above discussion by extending assumption [A1] to allow correlated random effects. We now focus on the marginal distribution of e_{it}^j, $f_{e^j}(e_{it}^j)$. Recall that it can be represented as $f_{e^j}(e_{it}^j) = \int f_{e^j | \alpha^j}(e_{it}^j | \alpha_i^j) f_{\alpha^j}(\alpha_i^j) d\alpha_i^j$ whether α_i^j and $\alpha_i^{j'}$ are correlated with each other or not. Therefore, the marginal pdf $f_{e^j}(e_{it}^j)$ can be evaluated by the simulated pdf $f_e^s(e_{it}) = \frac{1}{M} \sum_{m=1}^{M} f_\varepsilon(e_{it}^j - \alpha_{i(m)}^j) = \frac{1}{M} \sum_{m=1}^{M} f_\varepsilon \left(\varepsilon_{it(m)} \right)$.

According to Sklar's theorem, the joint pdf of $e_{it} = \left(e_{it}^1, \ldots, e_{it}^J \right)'$ can be represented as

$$f_e \left(e_{it} \right) = c \left(F_{e^1} \left(e_{it}^1 \right), \ldots, F_{e^J} \left(e_{it}^J \right); R \right) \times \prod_{j=1}^{J} f_{e^j} \left(e_{it}^j \right),$$

where the dependence parameter R captures the correlation between e_{it}^j's. Although that α_i^j and $\alpha_i^{j'}$ are correlated with each other, the following result holds for the conditional pdf of e_{it}^j

$$f_{e^j | \alpha} \left(e_{it}^j | \alpha_i^1, \ldots, \alpha_i^J \right) = f_{e^j} \left(e_{it}^j | \alpha_i^j \right),$$

which suggests that the information about $\alpha_i^{j'}$, where $j' \neq j$, is redundant for predicting e_{it}^j once we know α_i^j.

Given the information set $\alpha_i = (\alpha_i^1, \ldots, \alpha_i^J)$, the conditional joint pdf of $f_{e|\alpha}(e_{it}|\alpha_i)$ can be represented as

$$
\begin{aligned}
f_{e|\alpha}(e_{it}|\alpha_i) &= c\left(F_{e^1}\left(e_{it}^1|\alpha_i\right), \ldots, F_{e^J}\left(e_{it}^J|\alpha_i\right); R\right) \times \prod_{j=1}^{J} f_{e^j}\left(e_{it}^j|\alpha_i\right) \\
&= c\left(F_{e^1}\left(e_{it}^1|\alpha_i^j\right), \ldots, F_{e^J}\left(e_{it}^J|\alpha_i^J\right); R\right) \times \prod_{j=1}^{J} f_{e^j}\left(e_{it}^j|\alpha_i^j\right).
\end{aligned}
\tag{20}
$$

Since $e_{i1}, \ldots, e_{iT}$ are conditionally independent given α_i,

$$
\begin{aligned}
f_e(e_{i.}) &= \int f_{e|\alpha}(e_{i.}|\alpha_i) f_\alpha(\alpha_i) d\alpha_i \\
&= \int \prod_{t=1}^{T} f_{e|\alpha}(e_{it}|\alpha_i) f_\alpha(\alpha_i) d\alpha_i \\
&= \prod_{t=1}^{T} \int \left[c\left(F_{e^1}\left(e_{it}^1|\alpha_i^j\right), \ldots, F_{e^J}\left(e_{it}^J|\alpha_i^J\right); R\right) \right. \\
&\quad \left. \times \prod_{j=1}^{J} f_{e^j}\left(e_{it}^j|\alpha_i^j\right) \right] f_\alpha(\alpha_i) d\alpha_i,
\end{aligned}
\tag{21}
$$

where the third equality is due to (20). Thus, (21) can be evaluated by the simulated density

$$
\begin{aligned}
f_e^s(e_{i.}) &= \frac{1}{M}\sum_{m=1}^{M}\left\{ \prod_{t=1}^{T}\left[c\left(F_{e^1}\left(e_{it}^1|\alpha_{i(m)}^j\right), \ldots, F_{e^J}\left(e_{it}^J|\alpha_{i(m)}^J\right); R\right) \right. \right. \\
&\quad \left. \left. \times \prod_{j=1}^{J} f_{e^j}\left(e_{it}^j|\alpha_{i(m)}^j\right) \right] \right\}.
\end{aligned}
$$

Although $\alpha_{i(m)}^j$'s are drawn independently, their dependence will be captured by the copula. The logarithm of the simulated likelihood function of the whole sample is

$$
\ln L^s(\theta_1, \ldots, \theta_J, R) = \sum_{i=1}^{N} \ln f_e^s(e_{i.}),
\tag{22}
$$

where $f_e^s(e_{i.})$ is given by (21). Note that the simulated joint pdf $f_e^s(e_{i.})$ in (21) and the simulated joint pdf in (17) are equivalent. In other words, the model is estimated in the same way whether the random effects α_i's are correlated or not. However, it is worth mentioning that when the correlated random effects are correlated, the copula not only captures the dependence between ε_{it}^j and $\varepsilon_{it}^{j'}$, but also captures the dependence between α_i^j and $\alpha_i^{j'}$. Therefore, assumption [A1] can be further extended by allowing the possible correlation between the random effects, i.e., correlated α_i^j and $\alpha_i^{j'}$ for $j \neq j'$. The MSL estimator from (22) is also consistent.

4 Prediction of the inefficiency and technical efficiency

The prediction of the inefficiency and technical efficiency (TE) can follow the simulated approach proposed by Lai and Kumbhakar (2018). For a given division j, let $g(\cdot) : \mathbb{R}^+ \to \mathbb{R}$ be a continuous function of u_{it}^j and we are interested in the conditional expectation $\mathbb{E}\left(g(u_{it}^j)|e_{it}^j\right)$. For instance, if $g(u_{it}^j) = u_{it}^j$, then $\mathbb{E}\left(g(u_{it}^j)|e_{it}^j\right) = \mathbb{E}\left(u_{it}^j|e_{it}^j\right)$; if $g(u_{it}^j) = e^{-u_{it}^j}$, then $\mathbb{E}\left(g(u_{it}^j)|e_{it}^j\right) = \mathbb{E}\left(e^{-u_{it}^j}|\varepsilon_{it}^j\right)$.

The conditional expectation of the inefficiency given the composite error e_{it}^j is defined as

$$\mathbb{E}\left(u_{it}^j|e_{it}^j\right) = \int_0^\infty u_{it}^j f_{u^j|e^j}\left(u_{it}^j|e_{it}^j\right) du_{it}^j,$$

where

$$f_{u^j|e^j}\left(u_{it}^j|e_{it}^j\right) = \frac{\int_{-\infty}^\infty f_{u^j,e^j,\alpha^j}\left(u_{it}^j,e_{it}^j,\alpha_i^j\right) d\alpha_i^j}{f_{e^j}\left(e_{it}^j\right)}$$

$$= \frac{\int_{-\infty}^\infty f_{u^j|e^j,\alpha^j}\left(u_{it}^j|e_{it}^j,\alpha_i^j\right) f_{e^j|\alpha^j}\left(e_{it}^j|\alpha_i^j\right) f_{\alpha^j}\left(\alpha_i^j\right) d\alpha_i^j}{\int_{-\infty}^\infty f_{e^j|\alpha^j}\left(e_{it}^j|\alpha_i^j\right) f_{\alpha^j}\left(\alpha_i^j\right) d\alpha_i^j}.$$

Therefore, it follows that

$$\mathbb{E}\left(u_{it}^j|e_{it}^j\right) = \int_0^\infty u_{it}^j \int_{-\infty}^\infty f_{u^j|e^j,\alpha^j}\left(u_{it}^j|e_{it}^j,\alpha_i^j\right) \frac{f_{e^j|\alpha^j}\left(e_{it}^j|\alpha_i^j\right) f_{\alpha^j}\left(\alpha_i^j\right)}{\int_{-\infty}^\infty f_{e^j|\alpha^j}\left(e_{it}^j|\alpha_i^j\right) f_{\alpha^j}\left(\alpha_i^j\right) d\alpha_i^j} d\alpha_i^j du_{it}^j \tag{23a}$$

$$= \int_{-\infty}^\infty \left(\int_0^\infty u_{it}^j f_{u^j|e^j,\alpha^j}\left(u_{it}^j|e_{it}^j,\alpha_i^j\right) du_{it}^j\right) \frac{f_{e^j|\alpha^j}\left(e_{it}^j|\alpha_i^j\right) f_{\alpha^j}\left(\alpha_i^j\right)}{\int_{-\infty}^\infty f_{e^j|\alpha^j}\left(e_{it}^j|\alpha_i^j\right) f_{\alpha^j}\left(\alpha_i^j\right) d\alpha_i^j} d\alpha_i^j. \tag{23b}$$

Equation (23b) is related to the law of iterative expectation

$$\mathbb{E}\left(u_{it}^j|e_{it}^j\right) = \mathbb{E}\left[\mathbb{E}\left(u_{it}^j|e_{it}^j,\alpha_i^j\right)\Big| e_{it}^j\right]. \tag{24}$$

Given $f\left(u_{it}^j|e_{it}^j,\alpha_i^j\right) = f\left(u_{it}^j|\varepsilon_{it}^j\right)$ and the well-known result

$$\mathbb{E}\left(u_{it}^j|\varepsilon_{it}^j\right) = \widetilde{\mu}_{it}^j + \widetilde{\sigma}_{it}^j \left[\frac{\phi\left(-\widetilde{\mu}_{it}^j/\widetilde{\sigma}_{it}^j\right)}{1 - \Phi\left(-\widetilde{\mu}_{it}^j/\widetilde{\sigma}_{it}^j\right)}\right], \tag{25}$$

where $\varepsilon_{it}^j = v_{it}^j - u_{it}^j$, $\widetilde{\mu}_{it}^j = -\varepsilon_{it}^j \sigma_{uj,it}^2 / \left(\sigma_{uj,it}^2 + \sigma_{vj}^2 \right)$ and $\widetilde{\sigma}_{it}^{j2} = \sigma_{uj,it}^2 \sigma_{vj}^2 / \left(\sigma_{uj,it}^2 + \sigma_{vj}^2 \right)$, by assumptions [A2]–[A5], one may combine the results in Eqs. (24) and (25) to obtain the simulated estimator $\mathbb{E}^s \left(u_{it}^j | e_{it}^j \right)$ of $\mathbb{E} \left(u_{it}^j | e_{it}^j \right)$, which is defined as

$$\mathbb{E}^s \left(u_{it}^j | e_{it}^j \right) = \sum_{m=1}^{M} \left\{ \widetilde{\mu}_{it(m)}^j + \widetilde{\sigma}_{it}^j \left[\frac{\phi \left(-\widetilde{\mu}_{it(m)}^j / \widetilde{\sigma}_{it}^j \right)}{1 - \Phi \left(-\widetilde{\mu}_{it(m)}^j / \widetilde{\sigma}_{it}^j \right)} \right] \right\} W_{it(m)}^j, \quad (26)$$

where $\widetilde{\mu}_{it(m)}^j = -\varepsilon_{it(m)}^j \sigma_{uj,it}^2 / \left(\sigma_{uj,it}^2 + \sigma_{vj}^2 \right)$, $\varepsilon_{it(m)}^j = e_{it}^j - \alpha_{i(m)}^j$ and the weight $W_{it(m)}^j$ is defined as

$$W_{it(m)}^j = \frac{f_{\varepsilon^j}(e_{it}^j - \alpha_{i(m)}^j)}{\sum_{m=1}^{M} f_{\varepsilon^j}(e_{it}^j - \alpha_{i(m)}^j)} = \frac{f_{\varepsilon^j}(\varepsilon_{it(m)}^j)}{\sum_{m=1}^{M} f_{\varepsilon^j}(\varepsilon_{it(m)}^j)}. \quad (27)$$

Note that the weights $W_{it(m)}^j$ for $m = 1, \ldots, M$ have a sum equal to one. The estimator (26) has the form of the weighted average and gives the simulated estimator of the inefficiency $\mathbb{E} \left[u_{it}^j | e_{it}^j \right]$.

Moreover, given the result

$$\mathbb{E} \left(e^{-u_{it}^j} | \varepsilon_{it}^j \right) = \frac{1 - \Phi \left(\widetilde{\sigma}_{it} - \widetilde{\mu}_{it}^j / \widetilde{\sigma}_{it}^j \right)}{1 - \Phi \left(-\widetilde{\mu}_{it}^j / \widetilde{\sigma}_{it}^j \right)} \exp \left(-\widetilde{\mu}_{it}^j + \frac{1}{2} \widetilde{\sigma}_{it}^{j2} \right), \quad (28)$$

and

$$\mathbb{E} \left(e^{-u_{it}^j} | e_{it}^j \right) = \mathbb{E} \left[\mathbb{E} \left(e^{-u_{it}^j} | e_{it}^j, \alpha_i^j \right) \middle| e_{it}^j \right], \quad (29)$$

one can obtain the simulated estimator of TE,

$$\mathbb{E}^s \left(e^{-u_{it}^j} | e_{it}^j \right) = \sum_{m=1}^{M} \left[\frac{1 - \Phi \left(\widetilde{\sigma}_{it}^j - \widetilde{\mu}_{it(m)}^j / \widetilde{\sigma}_{it}^j \right)}{1 - \Phi \left(-\widetilde{\mu}_{it(m)}^j / \widetilde{\sigma}_{it}^j \right)} \exp \left(-\widetilde{\mu}_{it(m)}^j + \frac{1}{2} \widetilde{\sigma}_{it}^{j2} \right) \right] W_{it(m)}^j, \quad (30)$$

by the same approach used in (26).

In a manner similar to (26) and (30), we can also compute the conditional expectation of the marginal effect of $w_{it,k}^j$ on $\mathbb{E} \left(u_{it}^j | e_{it}^j \right)$. If $w_{it,k}^j$ is a continuous variable,

by the law of iterative expectation and the formula

$$\frac{\partial \mathbb{E}\left(u_{it}^j | \varepsilon_{it}^j\right)}{\partial w_{it,k}^j} = \delta_{j,k}\sigma_{uj,it}\left[\left(1 + \left(\frac{\widetilde{\mu}_{it}^j}{\widetilde{\sigma}_{it}^j}\right)^2\right)\left(\frac{\phi\left(-\widetilde{\mu}_{it}^j/\widetilde{\sigma}_{it}^j\right)}{1-\Phi\left(-\widetilde{\mu}_{it}^j/\widetilde{\sigma}_{it}^j\right)}\right) + \left(\frac{\widetilde{\mu}_{it}^j}{\widetilde{\sigma}_{it}^j}\right)\left(\frac{\phi\left(-\widetilde{\mu}_{it}^j/\widetilde{\sigma}_{it}^j\right)}{1-\Phi\left(-\widetilde{\mu}_{it}^j/\widetilde{\sigma}_{it}^j\right)}\right)^2\right] \quad (31)$$

one can obtain the simulated estimator

$$\frac{\partial \mathbb{E}^s\left(u_{it}^j | e_{it}^j\right)}{\partial w_{it,k}^j} = \sum_{m=1}^M \delta_{j,k}\sigma_{uj,it}\left[\left(1 + \left(\frac{\widetilde{\mu}_{it(m)}^j}{\widetilde{\sigma}_{it}^j}\right)^2\right)\left(\frac{\phi\left(-\widetilde{\mu}_{it(m)}^j/\widetilde{\sigma}_{it}^j\right)}{1-\Phi\left(-\widetilde{\mu}_{it(m)}^j/\widetilde{\sigma}_{it}^j\right)}\right) + \left(\frac{\widetilde{\mu}_{it(m)}^j}{\widetilde{\sigma}_{it}^j}\right)\left(\frac{\phi\left(-\widetilde{\mu}_{it(m)}^j/\widetilde{\sigma}_{it}^j\right)}{1-\Phi\left(-\widetilde{\mu}_{it(m)}^j/\widetilde{\sigma}_{it}^j\right)}\right)^2\right] W_{it(m)}^j.$$
$$(32)$$

Moreover, if $w_{it,k}^j$ is a dummy variable, the conditional expectation of the marginal effect of $w_{it,k}^j$ on $\mathbb{E}\left(u_{it}^j | e_{it}^j\right)$ is

$$\frac{\Delta\mathbb{E}\left(u_{it}^j | e_{it}^j\right)}{\Delta w_{it,k}^j} = \mathbb{E}\left(u_{it}^j | e_{it}^j, w_{it,k}^j = 1\right) - \mathbb{E}\left(u_{it}^j | e_{it}^j, w_{it,k}^j = 0\right), \quad (33)$$

where the estimators of $\mathbb{E}\left(u_{it}^j | e_{it}^j, w_{it,k}^j = 1\right)$ and $\mathbb{E}\left(u_{it}^j | e_{it}^j, w_{it,k}^j = 1\right)$ can be obtained using Eqs. (26) and (27) and substituting $w_{it,k}^j = 1$ and $w_{it,k}^j = 0$, respectively. Therefore, the inefficiency, TE, and marginal effect of $w_{it,k}^j$ on inefficiency can be estimated using (26), (30), (32) and (33), where the parameters are replaced by the MSL estimates obtained from (19).

5 Simulation

In this section, we conduct a Monte Carlo simulation to investigate the finite sample performance of the proposed simulated estimator. We consider the following data generating process (DGP) for the following system of two equations:

$$y_{it}^1 = \beta_0^1 + x_{it}^1\beta_1^1 + \alpha_i^1 + \varepsilon_{it}^1,$$
$$y_{it}^2 = \beta_0^2 + x_{it}^2\beta_1^2 + \alpha_i^2 + \varepsilon_{it}^2,$$

where $\varepsilon_{it}^j = v_{it}^j - u_{it}^j$, $v_{it}^j \sim i.i.d.N(0, \sigma_{vj}^2)$, and $u_{it}^j \sim N^+(0, \sigma_{uj,it}^2)$ for $j = 1, 2$. α_i^1 and α_i^2 are generated from a bivariate normal distribution with zero mean, correlation coefficient φ, and variance $\sigma_{\alpha j}^2$, where $j = 1, 2$, The heteroscedastic variance of u_{it}^j

is specified as $\sigma_{uj,it} = \exp(\delta_0^j + \delta_1^j w_{it}^j)$. The true parameters are set as

$$\beta_0^1 = 0.75, \beta_1^1 = 0.75, \sigma_{\alpha 1} = 0.1, \sigma_{v1} = 0.15, \delta_0^1 = -0.5, \delta_1^1 = 0.1,$$

and

$$\beta_0^2 = 1.5, \beta_1^2 = 0.5, \sigma_{\alpha 2} = 0.15, \sigma_{v2} = 0.1, \delta_0^2 = -0.75, \delta_1^2 = 0.5.$$

The exogenous variable x_{it}^1 is drawn from Uniform[5, 10], w_{it}^1 is drawn from $N(-1, 0.5^2)$, x_{it}^2 is drawn from Uniform[2, 5], and w_{it}^2 is drawn from $N(0, 0.4^2)$. Given the above marginal distributions of v_{it}^j and u_{it}^j, we generate the composite errors ε_{it}^1 and ε_{it}^2 from the Gaussian copula with the copula parameter ρ.[4] Under the above setting for $\delta's$, the sample means of $\sigma_{u1,it}$ and $\sigma_{u2,it}$ are about 0.549 and 0.472, and their standard deviations are about 0.029 and 0.028, respectively.

Below, we consider three experiments, labeled as Experiment I, Experiment II and Experiment III. In Experiment I, we intend to investigate the finite sample performance of the MSL estimator discussed in Sect. 3.2, and we also compare the MSL estimators with and without taking into account the dependence between divisions in estimation. In Experiment II, we intend to investigate the consequences of a misspecified copula. We examine the finite sample performance of the estimator when the random effects are correlated in Experiment III. Moreover, the total number of replications is 500 and 300 Halton draws are used in these experiments.

In Experiment I, we investigate the sampling patterns of our estimator by considering the following different combinations of N, T and ρ:

$$N = \{100, 200\}, \quad T = \{5, 10\} \text{ and } \rho = \{0.25, 0.75\}.$$

In the first experiment, we let the random effects be uncorrelated, i.e., we set $\varphi = 0$. In addition to estimating the model under the Gaussian copula specification, we also estimate the model using an independent copula, which ignores the correlation between the two equations, in order to compare the finite sample performances of these estimators. We summarize the biases and root mean squared errors (RMSE) in Tables 1 and 2, respectively. The results of independent and Gaussian copulas are given in Panels A and B of these tables.

From Tables 1 and 2, we find that the magnitudes of the biases and those of RMSEs are small for both copulas considered in these experiments. Nevertheless, we do not see any clear pattern that the biases of the Gaussian copula are smaller than those of the independent copula; but it can be observed from these tables that most of the biases in Panels A and B decrease as the sample size increases. However, Table 2 reveals that the RMSEs decrease when we increase either N or T for both independent and Gaussian copula specifications. Indeed, the RMSEs of the independent copula are larger than those of the Gaussian copula when N or T increases. In summary, the results in Tables 1 and 2 together provide evidence that the MSL estimators are consistent under both the

---

[4] More details about how to generate ε_{it}^1 and ε_{it}^2 from the Gaussian copula can be found in Lai and Huang (2013).

Table 1 Biases of the Monte Carlo experiments (DGP is Gaussian copula and α_i's are independent)

N	T	Equation 1						Equation 2						ρ
		β_0^1	β_1^1	$\sigma_{\alpha1}$	σ_{v1}	δ_0^1	δ_1^1	β_0^2	β_1^2	$\sigma_{\alpha2}$	σ_{v2}	δ_0^2	δ_1^2	
A. Independent copula														
$\rho = 0.25$														
100	5	0.0342	−0.0030	0.0198	−0.0011	−0.0205	−0.0101	−0.0011	−0.0032	0.0113	−0.0027	−0.0040	0.0075	N/A
100	10	0.0136	−0.0001	0.0187	0.0009	−0.0021	0.0034	−0.0115	−0.0001	0.0130	−0.0010	−0.0034	0.0116	N/A
200	5	0.0034	−0.0001	0.0130	0.0015	0.0107	0.0185	−0.0090	−0.0008	−0.0093	−0.0004	−0.0041	0.0027	N/A
200	10	0.0128	−0.0012	0.0134	−0.0005	−0.0028	−0.0026	−0.0137	0.0005	−0.0077	0.0007	−0.0033	0.0024	N/A
$\rho = 0.75$														
100	5	0.0248	−0.0019	0.0189	0.0008	−0.0312	−0.0173	−0.0026	−0.0029	0.0113	−0.0011	−0.0083	0.0096	N/A
100	10	0.0146	0.0001	0.0178	−0.0002	−0.0031	0.0002	−0.0157	0.0012	0.0138	−0.0007	−0.0005	0.0056	N/A
200	5	0.0054	−0.0003	0.0133	0.0000	0.0106	0.0145	−0.0081	−0.0008	−0.0097	−0.0010	−0.0024	0.0087	N/A
200	10	0.0131	−0.0013	0.0135	0.0001	−0.0028	−0.0006	−0.0115	0.0002	−0.0077	−0.0003	−0.0009	−0.0035	N/A
B. Gaussian copula														
$\rho = 0.25$														
100	5	−0.0250	0.0023	0.0017	−0.0002	−0.0279	−0.0175	0.0301	−0.0002	0.0080	−0.0012	−0.0111	0.0184	−0.0029
100	10	−0.0066	0.0001	0.0023	0.0003	−0.0063	−0.0003	0.0295	0.0009	0.0070	−0.0017	−0.0016	0.0051	−0.0014
200	5	−0.0016	0.0011	0.0009	0.0012	0.0043	0.0081	0.0166	−0.0016	0.0006	−0.0002	−0.0022	0.0069	0.0055
200	10	−0.0033	0.0014	0.0038	0.0007	0.0021	0.0064	0.0141	−0.0008	0.0016	0.0004	−0.0004	0.0051	0.0027
$\rho = 0.75$														
100	5	−0.0126	0.0010	0.0049	−0.0024	−0.0070	−0.0031	0.0347	−0.0010	0.0085	−0.0034	−0.0047	0.0030	0.0008
100	10	−0.0087	0.0006	0.0051	−0.0015	−0.0036	−0.0005	0.0298	0.0003	0.0094	−0.0005	−0.0052	0.0095	−0.0003
200	5	0.0032	0.0007	0.0007	0.0014	−0.0021	0.0007	0.0145	−0.0011	−0.0009	0.0014	−0.0037	0.0065	0.0040
200	10	0.0007	0.0010	0.0021	0.0008	−0.0027	0.0014	0.0120	−0.0005	−0.0007	0.0006	−0.0028	0.0010	−0.0001

The total number of replications is 500. N/A denotes not applicable

Table 2 RMSEs of the Monte Carlo experiments (DGP is Gaussian copula and $\alpha_i's$ are independent)

N	T	Equation 1						Equation 2						ρ
		β_0^1	β_1^1	$\sigma_{\alpha 1}$	$\sigma_{\upsilon 1}$	δ_0^1	δ_1^1	β_0^2	β_1^2	$\sigma_{\alpha 2}$	$\sigma_{\upsilon 2}$	δ_0^2	δ_1^2	
A. Independent copula														
$\rho = 0.25$														
100	5	0.0750	0.0095	0.0200	0.0271	0.0920	0.0694	0.0510	0.0137	0.0135	0.0286	0.0541	0.0925	N/A
100	10	0.0548	0.0070	0.0117	0.0154	0.0602	0.0484	0.0369	0.0097	0.0087	0.0130	0.0379	0.0595	N/A
200	5	0.0571	0.0071	0.0146	0.0179	0.0620	0.0484	0.0361	0.0096	0.0102	0.0161	0.0386	0.0675	N/A
200	10	0.0396	0.0050	0.0081	0.0113	0.0443	0.0351	0.0240	0.0062	0.0060	0.0098	0.0264	0.0447	N/A
$\rho = 0.75$														
100	5	0.0750	0.0095	0.0196	0.0249	0.0908	0.0656	0.0510	0.0137	0.0142	0.0245	0.0588	0.0903	N/A
100	10	0.0548	0.0068	0.0121	0.0164	0.0588	0.0475	0.0343	0.0090	0.0087	0.0129	0.0371	0.0627	N/A
200	5	0.0581	0.0071	0.0139	0.0179	0.0604	0.0516	0.0370	0.0097	0.0102	0.0166	0.0397	0.0579	N/A
200	10	0.0389	0.0051	0.0081	0.0111	0.0433	0.0333	0.0253	0.0065	0.0062	0.0092	0.0265	0.0443	N/A
B. Gaussian copula														
$\rho = 0.25$														
100	5	0.0793	0.0100	0.0231	0.0257	0.0841	0.0617	0.0526	0.0136	0.0130	0.0245	0.0572	0.0902	0.0461
100	10	0.0528	0.0068	0.0120	0.0147	0.0619	0.0482	0.0345	0.0090	0.0090	0.0128	0.0364	0.0616	0.0317
200	5	0.0501	0.0062	0.0152	0.0174	0.0630	0.0499	0.0346	0.0093	0.0088	0.0164	0.0387	0.0587	0.0307
200	10	0.0377	0.0047	0.0088	0.0112	0.0420	0.0343	0.0237	0.0063	0.0063	0.0094	0.0253	0.0406	0.0214
$\rho = 0.75$														
100	5	0.0582	0.0070	0.0182	0.0233	0.0736	0.0473	0.0400	0.0098	0.0103	0.0201	0.0532	0.0650	0.0214
100	10	0.0386	0.0047	0.0112	0.0141	0.0496	0.0330	0.0267	0.0064	0.0094	0.0117	0.0342	0.0454	0.0147
200	5	0.0394	0.0047	0.0118	0.0148	0.0520	0.0377	0.0270	0.0067	0.0069	0.0138	0.0349	0.0426	0.0144
200	10	0.0295	0.0036	0.0069	0.0096	0.0331	0.0243	0.0180	0.0044	0.0052	0.0081	0.0236	0.0292	0.0099

The total number of replications is 500. N/A denotes not applicable

independent and the Gaussian copula specifications, and the MSL estimator with the Gaussian copula is more efficient than that with the independent copula.

In our Experiment II, we reestimate the above model using Clayton, Gumbel, FGM (Farlie–Gumbel–Morgenstern) and AMH (Ali–Mikhail–Haq)[5] copulas for the sample size $N = 100$ and $T = 5$ in order to investigate how the MSL estimator performs under a misspecified copula. Both the Clayton and Gumbel copulas are Archimedean copulas, so their copulas can be represented in the form

$$C_{\text{Arch}}(\zeta_1, \zeta_2) = \psi^{-1} \left[\psi(\zeta_1) + \psi(\zeta_2) \right],$$

where $\psi : [0, 1] \to \mathbb{R}^+$ is called the Archimedean generator. $\psi(\cdot)$ is a continuous, strictly decreasing, and convex function and satisfies $\psi(0) = \infty$ and $\psi(1) = 0$. Moreover, the two arguments $\zeta_{j,it}$ are defined as $\zeta_{j,it} = F_{\varepsilon^j} \left(\varepsilon_{it}^j \right)$ for $j = 1, 2$ in our model. The corresponding density of the Archimedean copula is

$$c_{\text{Arch}}(\zeta_1, \zeta_2) = \frac{-\psi'' \left(C_{\text{Arch}}(\zeta_1, \zeta_2) \right) \psi'(\zeta_1) + \psi'(\zeta_2)}{\left[\psi' \left(C_{\text{Arch}}(\zeta_1, \zeta_2) \right) \right]^3}.$$

(i) If $\psi(\zeta) = \zeta^{-a} - 1$, we obtain the Clayton copula, which has the form

$$C_{\text{Clay}}(\zeta_1, \zeta_2) = \left(\zeta_1^{-a} + \zeta_2^{-a} - 1 \right)^{-1/a},$$

where $0 < a < \infty$ controls the strength of dependence. When $a = 0$, there is no dependence, and when $a = \infty$, there is prefect dependence.

(ii) If $\psi(\zeta) = (-\ln \zeta)^a$, we obtain the Gumbel copula, which has the form

$$C_{\text{Gum}}(\zeta_1, \zeta_2) = \exp \left\{ - \left[(-\ln \zeta_1)^a + (-\ln \zeta_2)^a \right]^{1/a} \right\},$$

where $a \geq 1$ controls the strength of dependence. When $a = 1$, there is no dependence, and when $a = \infty$, there is prefect dependence.

(iii) The third copula we considered is the FGM copula, which is defined as

$$C_{\text{FGM}}(\zeta_1, \zeta_2) = \zeta_1 \zeta_2 \left[1 + \kappa_F (1 - \zeta_1)(1 - \zeta_2) \right],$$

where $-1 \leq \kappa_F \leq 1$ is the copula parameter and ζ_1 and ζ_2 are defined as before. The corresponding copula density is

$$c_{\text{FGM}}(\zeta_1, \zeta_2) = 1 + \kappa_F (1 - 2\zeta_1)(1 - 2\zeta_2).$$

The Spearman's ρ of the FGM copula is $\kappa_F/3$, so it ranges between $-1/3$ and $1/3$.

---

[5] See Smith (2008) for more bivariate copulas.

(iv) The last copula is the AMH copula, whose copula function is defined as

$$C_{\text{AMH}}(\zeta_1, \zeta_2) = \frac{\zeta_1\zeta_2}{1 - \kappa_A(1 - \zeta_1)(1 - \zeta_2)},$$

where $-1 \leq \kappa_A \leq 1$ is the copula parameter and ζ_1 and ζ_2 are defined as before. The AMH copula density is

$$c_{\text{AMH}}(\zeta_1, \zeta_2)$$
$$= [1 + \kappa_A(\zeta_1\zeta_2 + \zeta_1 + \zeta_2 - 2 + \kappa_A(1 - \zeta_1)(1 - \zeta_2))][1 - \kappa_A(1 - \zeta_1)(1 - \zeta_2)]^{-3}.$$

In the DGP, the true copula is a Gaussian copula and we consider $\rho = 0.25$ and 0.75. We summarize biases and RMSEs of our estimates in Table 3. Since the dependence parameters of the Clayton, Gumbel, FGM and AMH copulas are not directly comparable with the Gaussian copula parameter ρ, we do not report the biases and RMSEs of the dependence parameters. We found that the bias of β_0^1 is relatively large when the Gumbel copula is used and $\rho = 0.25$. However, the biases of the estimates of the other parameters are quite close to the estimates of the Gaussian copula in Tables 1 and 2. Most of the estimates have biases slightly larger than the estimates from the Gaussian copula, but they are also of a small magnitude. Moreover, there is no clear evidence showing that the estimators of the misspecified copulas have larger biases than that of the independent copula. For the RMSEs, we found that some RMSEs from the misspecified copulas are slightly larger than those from the Gaussian copula, but there is no clear pattern. On the other hand, it is worth mentioning that the RMSEs of the independent copula are in general larger than the other copulas, which suggests that using a copula to capture the dependence between the divisions is helpful for improving the estimation efficiency even if the copula is misspecified. From the above experiment, we may conclude that if a misspecified copula is used, the ML estimator may have a slightly larger bias and RMSE than that from the correctly specified copula, but the problem may not be serious as one expects.

In Experiment III, we allow correlation between the random effects, i.e., α_i^1 and α_i^2 are correlated with correlation coefficient φ. Similar to Tables 1 and 2, we consider independent and Gaussian copulas for $(N, T) = \{(100, 5), (200, 10)\}$ and $\rho = \{0.25, 0.75\}$. To investigate how the degree of the correlation between the random effects affect the performance of the MSL estimator, we set $\varphi = \{0.25, 0.75\}$ and compare their differences. We summarize the simulation results in Tables 4 and 5. It is worth mentioning that the same model specification is used in Experiments I and III, but their DGPs are slightly different. The only difference between them is the value of φ. $\varphi = 0$ in Experiment I. Similar to what we have observed from Table 1, the biases from the independent copula are quite close to that of the Gaussian copula. This is because both independent and Gaussian copulas give consistent estimators whether α_i^j's are correlated or not. However, it is clear that the estimator using the Gaussian copula is more efficient than the estimator using the independent copula in terms of RMSE. By comparing the biases in Tables 1 and 5, we find that the biases of ρ in Table 5 are slightly larger than those in Table 1 because the correlation between the

Table 3 Bias and RMSE when different copulas are used ($N = 100$, $T = 5$, DGP is Gaussian copula and α_i's are independent)

Copula	ρ	Equation 1						Equation 2					
		β_0^1	β_1^1	$\sigma_{\alpha 1}$	σ_{v1}	δ_0^1	δ_1^1	β_0^2	β_1^2	$\sigma_{\alpha 2}$	σ_{v2}	δ_0^2	δ_1^2
Bias													
Gaussian	0.25	− 0.0250	0.0023	0.0017	− 0.0002	− 0.0279	− 0.0175	0.0301	− 0.0002	0.0080	− 0.0012	− 0.0111	0.0184
Indep.	0.25	0.0342	− 0.0030	0.0198	− 0.0011	− 0.0205	− 0.0101	− 0.0011	− 0.0032	0.0113	− 0.0027	− 0.0040	0.0075
Clayton	0.25	− 0.0205	0.0022	0.0044	− 0.0048	− 0.0149	− 0.0119	0.0299	0.0003	0.0078	− 0.0029	− 0.0075	0.0161
Gumbel	0.25	− 0.2478	0.0022	0.0053	0.0028	− 0.0308	− 0.0112	0.0270	0.0002	0.0086	0.0022	− 0.0195	0.0175
FGM	0.25	− 0.0215	0.0020	0.0031	− 0.0020	− 0.0189	− 0.0096	0.0301	− 0.0005	0.0074	0.0001	− 0.0128	0.0204
AMH	0.25	0.0390	0.0021	− 0.0003	− 0.0127	0.0058	− 0.0182	0.0734	0.0005	0.0081	− 0.0116	0.0113	0.0033
Gaussian	0.75	− 0.0126	0.0010	0.0049	− 0.0024	− 0.0070	− 0.0031	0.0347	− 0.0010	0.0085	− 0.0034	− 0.0047	0.0030
Indep.	0.75	0.0248	− 0.0019	0.0189	0.0008	− 0.0312	− 0.0173	− 0.0026	− 0.0029	0.0113	− 0.0011	− 0.0083	0.0096
Clayton	0.75	0.0126	0.0017	0.0222	− 0.0367	0.0556	− 0.0088	0.0524	− 0.0001	0.0130	− 0.0273	0.0502	0.0024
Gumbel	0.75	− 0.0372	0.0012	0.0054	0.0330	− 0.0661	0.0039	0.0195	− 0.0007	0.0113	0.0230	− 0.0509	0.0209
FGM	0.75	− 0.0375	0.0016	0.0071	− 0.0022	− 0.0596	− 0.0054	0.0153	− 0.0001	0.0085	− 0.0013	− 0.0546	0.0203
AMH	0.75	0.0613	0.0016	− 0.0027	− 0.0039	0.0285	− 0.0138	0.0808	− 0.0008	0.0082	− 0.0091	0.0342	− 0.0074
RMSE													
Gaussian	0.25	0.0793	0.0100	0.0231	0.0257	0.0841	0.0617	0.0526	0.0136	0.0130	0.0245	0.0572	0.0902
Indep.	0.25	0.0750	0.0095	0.0200	0.0271	0.0920	0.0694	0.0510	0.0137	0.0135	0.0286	0.0541	0.0925
Clayton	0.25	0.0772	0.0094	0.0212	0.0265	0.0891	0.0658	0.0517	0.0133	0.0135	0.0245	0.0599	0.0897
Gumbel	0.25	0.0095	0.0095	0.0214	0.0273	0.0907	0.0655	0.0525	0.0133	0.0129	0.0272	0.0593	0.0906

Table 3 continued

Copula	ρ	Equation 1						Equation 2					
		β_0^1	β_1^1	$\sigma_{\alpha 1}$	σ_{v1}	δ_0^1	δ_1^1	β_0^2	β_1^2	$\sigma_{\alpha 2}$	σ_{v2}	δ_0^2	δ_1^2
FGM	0.25	0.0755	0.0094	0.0221	0.0260	0.0916	0.0681	0.0519	0.0132	0.0130	0.0251	0.0580	0.0862
AMH	0.25	0.0819	0.0104	0.0252	0.0304	0.0826	0.0625	0.0593	0.0156	0.0146	0.0279	0.0512	0.0848
Gaussian	0.75	0.0582	0.0070	0.0182	0.0233	0.0736	0.0473	0.0400	0.0098	0.0103	0.0201	0.0532	0.0214
Indep.	0.75	0.0750	0.0095	0.0196	0.0249	0.0908	0.0656	0.0510	0.0137	0.0142	0.0245	0.0588	0.0903
Clayton	0.75	0.0639	0.0078	0.0163	0.0235	0.0709	0.0474	0.0420	0.0108	0.0124	0.0199	0.0446	0.0687
Gumbel	0.75	0.0606	0.0072	0.0172	0.0249	0.0821	0.0534	0.0442	0.0103	0.0109	0.0278	0.0724	0.0706
FGM	0.75	0.0647	0.0079	0.0189	0.0239	0.0837	0.0636	0.0486	0.0122	0.0124	0.0215	0.0571	0.0787
AMH	0.75	0.0925	0.0120	0.0267	0.0362	0.0846	0.0656	0.0650	0.0169	0.0164	0.0337	0.0555	0.0858

The total number of replications is 500

Table 4 Bias when α_i's are correlated (DGP is Gaussian copula)

N	T	φ	Equation 1						Equation 2						ρ
			β_0^1	β_1^1	$\sigma_{\alpha 1}$	σ_{v1}	δ_0^1	δ_1^1	β_0^2	β_1^2	$\sigma_{\alpha 2}$	σ_{v2}	δ_0^2	δ_1^2	
Independent copula															
$\rho = 0.25$															
100	5	0.25	−0.0268	0.0026	0.0017	−0.0007	−0.0285	−0.0184	0.0204	0.0015	0.0044	0.0007	−0.0121	0.0187	N/A
200	10	0.25	−0.0044	0.0018	0.0034	0.0003	0.0027	0.0048	0.0179	−0.0005	0.0077	−0.0006	−0.0002	−0.0002	N/A
100	5	0.75	−0.0237	0.0022	−0.0005	0.0008	−0.0229	−0.0117	0.0197	0.0016	0.0049	0.0010	−0.0143	0.0134	N/A
200	10	0.75	−0.0047	0.0018	0.0047	0.0002	−0.0019	0.0016	0.0178	−0.0002	0.0117	0.0007	−0.0036	0.0009	N/A
$\rho = 0.75$															
100	5	0.25	−0.0263	0.0025	−0.0017	0.0031	−0.0277	−0.0154	−0.0013	0.0037	−0.0005	0.0015	−0.0154	0.0141	N/A
200	10	0.25	−0.0065	0.0020	0.0041	0.0007	−0.0005	0.0024	0.0190	−0.0005	0.0126	0.0008	−0.0034	0.0028	N/A
100	5	0.75	−0.0270	0.0032	0.0018	−0.0037	−0.0135	−0.0105	0.0173	0.0026	0.0047	0.0002	−0.0142	0.0185	N/A
200	10	0.75	−0.0077	0.0021	0.0035	0.0002	−0.0022	0.0016	0.0167	−0.0003	0.0078	−0.0005	−0.0011	0.0045	N/A
Gaussian copula															
$\rho = 0.25$															
100	5	0.25	−0.0233	0.0023	−0.0043	0.0020	−0.0208	−0.0098	0.0264	0.0002	0.0011	0.0005	−0.0102	0.0157	0.0102
200	10	0.25	−0.0059	0.0019	0.0009	0.0013	0.0008	0.0018	0.0154	−0.0003	0.0041	0.0013	−0.0030	0.0042	0.0050
100	5	0.75	−0.0238	0.0023	−0.0177	0.0066	−0.0153	−0.0068	0.0006	0.0037	−0.0132	0.0078	−0.0117	0.0187	0.0391
200	10	0.75	−0.0054	0.0019	−0.0037	0.0029	0.0002	0.0025	0.0171	0.0000	0.0039	0.0036	−0.0012	0.0052	0.0146
$\rho = 0.75$															
100	5	0.25	−0.0126	0.0010	−0.0087	0.0032	−0.0103	−0.0031	0.0250	0.0006	−0.0029	0.0042	−0.1035	0.0136	0.0054
200	10	0.25	0.0016	0.0009	−0.0058	0.0032	−0.0081	0.0012	0.0153	−0.0007	−0.0033	0.0040	−0.0037	0.0033	0.0023
100	5	0.75	−0.0166	0.0019	−0.0638	0.0191	0.0041	−0.0014	0.0113	0.0014	−0.0391	0.0221	−0.0019	0.0116	0.0208
200	10	0.75	0.0037	0.0008	0.0291	0.0108	0.0035	0.0017	0.0182	0.0021	−0.0200	0.0114	0.0011	0.0007	0.0098

The total number of replications is 500. N/A denotes not applicable. φ is the correlation coefficient of α_i^1 and α_i^2

random effects has been picked up by the copula parameter in the model. Moreover, we also find that the estimator using the independent copula loses efficiency as the degree of dependence between ε_{it}^j's (or α_i^j's) increases.

In summary, we find that using either the independent copula or Gaussian copula gives a consistent estimator of the model parameter. Misspecification of a copula may cause estimation bias, but the problem may not be too serious. Even though we do not estimate the correlation coefficient φ between the random effects in the model, the MSL estimator is still consistent and the correlation will be automatically captured by the copula parameter, maintaining the estimation efficiency.

6 The empirical application

This paper applies the proposed method to the production frontier of Taiwan's international hotels; each hotel has two divisions, accommodation and restaurant, i.e., $J = 2$ for the system in (1). The data are derived from the annual report of the Taiwan Tourism Bureau at the Ministry of Transportation and Communications. Our sample is an unbalanced panel data, which contains 725 sample observations from 61 international grand hotels during 2001–2013. The minimum observed time period is 6 years, while the maximum period is 13 years.

For the accommodation division, the output is measured in total revenue (y_1), while the inputs include the total number of workers (x_{11}), the total number of rooms (x_{12}), and other expenses (x_{13}), which include utilities, materials, maintenance fees and so on. The output and inputs are allocable within the accommodation division. The output for the restaurant division is also measured by the total revenue (y_2), while the corresponding inputs are the total number of workers (x_{21}), the floor area of the restaurant (x_{22}) and other expenses (x_{23}), including utilities, materials and so on. Again, the output and inputs are attributable and accountable within the restaurant division. All revenues and other expenses are measured in New Taiwan dollars (NT$). Logarithms are applied to outputs and inputs. The exogenous determinants of the inefficiencies of the two divisions include the scale of the hotel (w_1) and the area dummy variable (w_2). The scale variable w_1 ranges from 1 to 5.[6] The dummy variable w_2 is equal to one if the hotel is located in a scenic area and zero otherwise. Since these two divisions of a hotel share certain common characteristics, such as the same DMU, brand and location, we expect the two outputs y_1 and y_2 should be correlated with each other and the composite errors to be also.

The empirical model is specified as

$$y_{it}^j = \beta_0^j + \beta_1^j x_{1,it}^j + \beta_2^j x_{2,it}^j + \beta_3^j x_{3,it}^j + \alpha_i^j + \varepsilon_{it}^j, \quad j = 1, 2. \tag{34}$$

---

[6] $w_1 = 1$ if the total number of employees is less than 100; $w_1 = 2$ if the total number of employees is larger than 100 and less than 200; $w_1 = 3$ if the total number of employees is larger than 200 and less than 300; and so on. $w_1 = 5$ if the total number of employees is more than 400.

Table 5 RMSE when α_i's are correlated (DGP is Gaussian copula)

N	T	φ	Equation 1						Equation 2						ρ
			β_0^1	β_1^1	$\sigma_{\alpha 1}$	σ_{v1}	δ_0^1	δ_1^1	β_0^2	β_1^2	$\sigma_{\alpha 2}$	σ_{v2}	δ_0^2	δ_1^2	
Independent copula															
$\rho = 0.25$															
100	5	0.25	0.0820	0.0103	0.0253	0.0266	0.0903	0.0719	0.0563	0.0143	0.0144	0.0241	0.0579	0.0897	N/A
200	10	0.25	0.0396	0.0050	0.0091	0.0109	0.0449	0.0351	0.0249	0.0065	0.0066	0.0094	0.0261	0.0435	N/A
100	5	0.75	0.0821	0.0104	0.0236	0.0261	0.0864	0.0669	0.0543	0.0142	0.0132	0.0238	0.0605	0.0943	N/A
200	10	0.75	0.0396	0.0050	0.0082	0.0110	0.0447	0.0352	0.0261	0.0068	0.0067	0.0098	0.0266	0.0423	N/A
$\rho = 0.75$															
100	5	0.25	0.0805	0.0103	0.0254	0.0250	0.0880	0.0678	0.0570	0.0147	0.0132	0.0234	0.0580	0.0888	N/A
200	10	0.25	0.0383	0.0049	0.0099	0.0114	0.0449	0.0361	0.0244	0.0063	0.0075	0.0093	0.0256	0.0421	N/A
100	5	0.75	0.0783	0.0099	0.0220	0.0258	0.0837	0.0670	0.0540	0.0137	0.0136	0.0238	0.0572	0.0942	N/A
200	10	0.75	0.0396	0.0045	0.0090	0.0114	0.0446	0.0350	0.0251	0.0068	0.0069	0.0095	0.0249	0.0436	N/A
Gaussian copula															
$\rho = 0.25$															
100	5	0.25	0.0762	0.0097	0.0255	0.0250	0.0900	0.0653	0.0509	0.0132	0.0129	0.0240	0.0559	0.0861	0.0451
200	10	0.25	0.0367	0.0047	0.0086	0.0111	0.0425	0.0347	0.0237	0.0062	0.0062	0.0089	0.0250	0.0405	0.0226
100	5	0.75	0.0761	0.0094	0.0280	0.0251	0.0873	0.0660	0.0533	0.0135	0.0145	0.0237	0.0560	0.0886	0.0442
200	10	0.75	0.0370	0.0047	0.0098	0.0120	0.0440	0.0326	0.0239	0.0062	0.0069	0.0094	0.0257	0.0419	0.0225
$\rho = 0.75$															
100	5	0.25	0.0588	0.0072	0.0178	0.0225	0.0723	0.0485	0.0398	0.0097	0.0095	0.0209	0.0540	0.0603	0.0211
200	10	0.25	0.0269	0.0032	0.0077	0.0097	0.0354	0.0254	0.0187	0.0046	0.0053	0.0080	0.0241	0.0289	0.0099
100	5	0.75	0.0553	0.0067	0.0281	0.0211	0.0676	0.0454	0.0408	0.0098	0.0100	0.0191	0.0540	0.0624	0.0194
200	10	0.75	0.0283	0.0033	0.0107	0.0097	0.0346	0.0240	0.0185	0.0046	0.0062	0.0081	0.0250	0.0294	0.0107

The total number of replications is 500. N/A denotes not applicable. φ is the correlation coefficient of α_i^1 and α_i^2

Table 6 Empirical results

	I. Separate Estimation		II. Joint estimation Independent copula		III. Joint estimation Gaussian copula	
	coef.	s.e.[a]	coef.	s.e.	coef.	s.e.
Accommodation division						
β_1^1	0.3104	0.0086***[b]	0.3426	0.0214***	0.3000	0.0003***
β_2^1	0.4276	0.0117***	0.5229	0.0183***	0.3958	0.0011***
β_3^1	0.2835	0.0044***	0.2929	0.0040***	0.4411	0.0012***
β_0^1	13.0459	0.0583***	12.3334	0.0376***	11.9497	0.0029***
Random components						
$\sigma_{\alpha 1}$	0.2808	0.0032***	0.3309	0.0034***	0.3244	0.0006***
σ_{v1}	0.0766	0.0016***	0.0758	0.0018***	0.0021	0.0004***
$\sigma_{u1} = \exp(\delta_1^1 w_1 + \delta_2^1 w_2 + \delta_0^1)$						
δ_1^1	−0.0863	0.0093***	−0.0872	0.0080***	−0.0759	0.0033***
δ_2^1	−0.0304	0.0220***	−0.0583	0.0184***	−0.0549	0.0117***
δ_0^1	−1.0901	0.0344***	−1.0606	0.0299***	−0.8321	0.0127***
Restaurant division						
β_1^2	0.4295	0.0083***	0.4611	0.0087***	0.3778	0.0044***
β_2^2	0.0442	0.0021***	0.0451	0.0018***	0.0223	0.0012***
β_3^2	0.1627	0.0049***	0.1566	0.0046***	0.3883	0.0043***
β_0^2	15.1821	0.0815***	15.104	0.0704***	13.8552	0.0340***
Random components						
$\sigma_{\alpha 2}$	0.4038	0.0075***	0.4855	0.0096***	0.3063	0.0027***
σ_{v2}	0.1373	0.0018***	0.1377	0.0018***	0.1675	0.0022***
$\sigma_{u2} = \exp(\delta_1^2 w_1 + \delta_2^2 w_2 + \delta_0^2)$						
δ_1^2	−1.9918	0.0759***	−1.9702	0.0907***	−0.9068	0.0315***
δ_2^2	−1.8949	0.0930***	−1.8748	0.0940***	−1.0208	0.0416***
δ_0^2	2.1571	0.1309***	2.1193	0.1588***	0.5395	0.0613***
Copula coeff.						
ρ	N/A[c]		N/A		0.8431	0.0068***
$\ln L$	362.0287		356.0293		581.9081	

[a] All reported standard errors are computed using the sandwich formula
[b] ***, ** and * represent significance at the 1%, 5% and 10% levels
[c] N/A denotes not applicable

To obtain the likelihood function of the empirical model, we differentiate Eq. (8) with respect to ε_{it}^1 and ε_{it}^2 and obtain the joint pdf

$$f_\varepsilon\left(\varepsilon_{it}^1, \varepsilon_{it}^2\right) = c\left(F_{\varepsilon^1}\left(\varepsilon_{it}^1\right), F_{\varepsilon^2}\left(\varepsilon_{it}^2\right)\right) f_{\varepsilon^1}\left(\varepsilon_{it}^1\right) f_{\varepsilon^2}\left(\varepsilon_{it}^2\right), \tag{35}$$

Table 7 Predicted Inefficiencies, TEs and marginal effects

	Accommodation		Restaurant	
	mean	s.d.	mean	s.d.
Panel A. Inefficiency $\mathbb{E}\left(u_{it}^j \mid \varepsilon_{it}^j\right)$				
I. Separate estimation	0.2071	0.0897	0.0874	0.2795
II. Independent copula	0.2130	0.0701	0.0848	0.2795
III. Gaussian copula	0.2506	0.0886	0.1301	0.2023
Panel B. TE $\mathbb{E}\left(e^{-u_{it}^j} \mid \varepsilon_{it}^j\right)$				
I. Separate estimation	0.8231	0.0667	0.9356	0.1278
II. Independent copula	0.8181	0.0523	0.9404	0.1243
III. Gaussian copula	0.7914	0.0629	0.8949	0.1171
Panel C. Marginal effects				
$10 \times \partial \mathbb{E}\left(u_{it}^j \mid \varepsilon_{it}^j\right)/\partial w_1$				
I. Separate estimation	-0.0681	0.0305	-0.6842	1.0502
II. Independent copula	-0.0660	0.0243	-0.6908	1.0213
III. Gaussian copula	-0.0002	0.0006	-0.6459	0.5155
$10 \times \Delta \mathbb{E}\left(u_{it}^j \mid \varepsilon_{it}^j\right)/\Delta w_2$				
I. Separate estimation	-0.0469	0.0307	-0.0520	0.0214
II. Independent copula	-0.0952	0.0487	-0.8343	2.0495
III. Gaussian copula	-0.1012	0.0404	-0.7774	0.7399

where $c(F_{\varepsilon^1}(\varepsilon_{it}^1), F_{\varepsilon^2}(\varepsilon_{it}^2)) = \partial C(F_{\varepsilon^1}(\varepsilon_{it}^1), F_{\varepsilon^2}(\varepsilon_{it}^2))/\partial F_{\varepsilon^1}(\varepsilon_{it}^1)\partial F_{\varepsilon^2}(\varepsilon_{it}^2)$ is the copula density.

For the sake of comparison, we consider three different model specifications: Model (I) is the separate estimation, where each division is estimated by a single-equation stochastic frontier model with random effects. Model (II) uses the joint estimation, where the joint pdf of (35) and an independent copula are used. For the independent copula, we have $c(F_{\varepsilon^1}(\varepsilon_{it}^1), F_{\varepsilon^2}(\varepsilon_{it}^2)) = 1$ and, therefore,

$$f_\varepsilon(\varepsilon_{it}^1, \varepsilon_{it}^2) = f_{\varepsilon^1}\left(\varepsilon_{it}^1\right) f_{\varepsilon^2}\left(\varepsilon_{it}^2\right). \tag{36}$$

The independent copula implies that the two divisions are uncorrelated, so the correlation between the two divisions is ignored in the empirical model. It is worth mentioning that the simulated likelihood functions of Models (I) and (II) are different since their simulated likelihood functions are evaluated in different ways. In Model (I), the log-likelihood function of the system is the sum of the logarithms of the simulated likelihood functions of each single equation. For Model (II), it is the logarithm of the simulated product of the marginal pdfs. We may expect that both of them give consistent estimators of the parameters as long as the marginal probability model is correctly specified. Model (III) uses the joint pdf in (35) and the Gaussian copula,

where the correlation between the two divisions is captured by the copula parameter ρ. Under the Gaussian copula assumption, the joint pdf of ε_{it}^1 and ε_{it}^2 is

$$f_\varepsilon \left(\varepsilon_{it}^1, \varepsilon_{it}^2 \right) = c \left(F_{\varepsilon^1} \left(\varepsilon_{it}^1 \right), F_{\varepsilon^2} \left(\varepsilon_{it}^2 \right) \right) f_{\varepsilon^1} \left(\varepsilon_{it}^1 \right) f_{\varepsilon^2} \left(\varepsilon_{it}^2 \right), \qquad (37)$$

where

$$c \left(\varsigma_{it}^1, \varsigma_{it}^2 \right) = \frac{1}{\sqrt{1-\rho^2}} \exp \left(\frac{\left(\varsigma_{it}^1\right)^2 + \left(\varsigma_{it}^2\right)^2}{2} + \frac{2\rho\varsigma_{it}^1\varsigma_{it}^2 - \left(\varsigma_{it}^1\right)^2 - \left(\varsigma_{it}^2\right)^2}{2\left(1-\rho^2\right)} \right), \qquad (38)$$

$\varsigma_{it}^j = \Phi^{-1}\left(r_{it}^j \right) = \Phi^{-1}\left(F_{\varepsilon^j}\left(\varepsilon_{it}^j \right) \right)$ and $r_{it}^j = F_{\varepsilon^j}(\varepsilon_{it}^j)$ for $j = 1, 2$. Under the Gaussian copula, the linear correlation between F_{ε^1} and F_{ε^2} is

$$\gamma_{\text{Spearman}} = \frac{6}{\pi} \arcsin \frac{\rho}{2}, \qquad (39)$$

which measures the correlation of the two divisions in terms of the cdfs of ε_{it}^1 and ε_{it}^2 and is also called the Spearman's rank correlation coefficient of ε_{it}^1 and ε_{it}^2.

We summarize the simulated ML estimates of Models (I), (II) and (III) in Table 6. As we expected, the estimated parameters of Models (I) and (II) are quite close to each other since they have the same probability model but their likelihood functions are evaluated in different ways. The estimates of Model (III) are slightly different from those of Models (I) and (II). Model (III) allows for correlation between the two divisions, and thus, its MSL estimator is more efficient. For the accommodation division, the input elasticities of x_{11} and x_{12} from the separate model and the independent copula model are smaller than those from the Gaussian copula models, while the elasticity of x_{13} from the Gaussian copula model is larger than the other two specification. A similar pattern can also be found in the restaurant division. Moreover, almost all standard errors of Model (III) are much smaller than those obtained from Models (I) and (II). The coefficients of the exogenous determinants of inefficiencies have the same signs in the three models. Given the estimated Gaussian copula parameter $\widehat{\rho} = 0.8431$, the corresponding Spearman's rank coefficient $\widehat{\gamma}_{\text{Spearman}}$ is 0.8311, which suggests that $F_{\varepsilon^1}(\varepsilon_{it}^1)$ and $F_{\varepsilon^2}(\varepsilon_{it}^2)$ are highly correlated and that the two divisions are also. Both Models (I) and (II) give less efficient estimators.

Once we obtain the estimated parameters, we may predict the inefficiencies, TEs and the marginal effects of the exogenous determinants (w_{it}) on the inefficiency using (26), (30), (32) and (33). Their sample statistics are summarized in Panels A, B and C of Table 7, respectively. Given the estimators of Models (I) and (II), it looks like the inefficiencies of the two divisions from Models (I) and (II) are lower than those from the Gaussian copula, and, therefore, the TEs are larger in Models (I) and (II) than those in Model (III). The predicted marginal effects from the three models consistently suggest that increasing the scale of a hotel has negative effects on inefficiencies of the accommodation and restaurant sectors and is thus helpful in improving the technical efficiency. Moreover, our results also show that the hotels located in scenic areas are also less efficient compared with those located in cities.

7 Conclusion

In this paper, we used a copula-based simulated maximum likelihood approach to estimate multiple panel stochastic frontier regressions with correlated composite errors. The innovation of the proposed model is to address the unobserved heterogeneity of the panel data. Compared with the separate estimation, the joint estimation of the multiple SF regressions is more efficient since the joint approach takes into consideration the correlation among composite errors. Therefore, this paper provides a better prediction of the inefficiency under this panel data framework. Although the system model considered in the current paper focuses on a production frontier system, it is a straightforward extension to modify the current production frontier system to a cost frontier system or to a system consisting of a combination of the production and cost frontiers.

References

Aigner DJ, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontier production models. J Econom 6:21–37

Amsler C, Schmidt P (2021) A Survey of the use of copulas in stochastic frontier models. In: Parmeter CF, Sickles RC (eds) Advances in efficiency and productivity analysis. NAPW 2018. Springer Proceedings in Business and Economics. Springer, Cham. https://doi.org/10.1007/978-3-030-47106-4_6

Amsler C, Prokhorov A, Schmidt P (2014) Using copulas to model time dependence in stochastic frontier models. Econom Rev 33(5–6):497–522

Cherubini U, Luciano E, Vecchiato W (2004) Copula methods in finance. Wiley, New York

Genius M, Stefanou S, Tzouvelekas V (2012) Measuring productivity growth under factor non-substitution: an application to US steam-electric power generation utilities. Eur J Oper Res 220:844–852

Greene WH (2003) Simulated likelihood estimation of the normal-gamma stochastic frontier function. J Prod Anal 19:179–190

Huang T-H, Chiang D-L, Chao S-W (2017a) A new approach to jointly estimating the lerner index and cost efficiency for multi-output banks under a stochastic meta-frontier framework. Q Rev Econ Finance 65:212–226

Huang T-H, Lin C-I, Chen K-C (2017b) Evaluating efficiencies of Chinese commercial banks in the context of stochastic multistage technologies. Pac Basin Finance J 41:93–110

Huang T-H, Chen K-C, Lin C-I (2018a) An extension from network DEA to copula-based network SFA: evidence from the U.S. commercial banks in 2009. Q Rev Econ Finance 67:51–62

Huang T-H, Liu N-H, Kumbhakar SC (2018b) Joint estimation of the Lerner index and cost efficiency using copula methods. Empir Econ 54:799–822

Lai H-P, Huang CJ (2013) Maximum likelihood estimation of seemingly unrelated stochastic frontier regressions. J Prod Anal 40(1):1–14

Lai H-P, Kumbhakar SC (2018) Panel data stochastic frontier model with determinants of persistent and transient inefficiency. Eur J Oper Res 271(2):746–755

Mundlak Y (1978) On the pooling of time series and cross section data. Econometrica 46:69–85

Schweizer B, Sklar A (1983) Probability metric spaces. Elsevier, New York

Sklar A (1959) Fonctions de Répartition à n Dimensions et Leurs Marges. Publications de l'Institut de Statistique de L'Université de Paris 8:229–231

Smith MD (2008) Stochastic frontier models with dependent error components. Econ J 11:172–192

Trivedi PK, Zimmer DM (2005) Copula modeling: an introduction for practitioners. Found Trends Econom 1(1):1–111